JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 162, 71–91 (1991)

Convergence in Distribution of Products of $d \times d$ Random Matrices

CHI-CHANG LO AND ARUNAVA MUKHERJEA

Department of Mathematics, University of South Florida, Tampa, Florida 33620-5700

Submitted by David Aldous

Received November 17, 1989

In this paper we present limit theorems on the convergence in distribution of products of i.i.d. matrices based on easily verifiable conditions on the support of the distribution.

1. INTRODUCTION

In recent years, there have been a number of studies on convergence in distribution of products of random matrices (see [14, 6, 8, 1, 4, 11, 5]). Some of these studies are quite general and involve real matrices. However, there remain many open questions and many problems remain unsolved even in the much simpler case of bistochastic matrices (i.e., nonnegatives matrices with each row and column sum one). For example, if A_1, A_2, ..., is a sequence of i.i.d. (that is, independent and identically distributed) $d \times d$ bistochastic matrices (with usual topology), then how can we decide in terms of *easily verifiable* conditions on the support of the distribution of A_i whether the sequence of products $M_n = A_1 A_2 \cdots A_n$ converges in distribution to some limiting distribution and how can we identify the limit? Note that $d \times d$ stochastic matrices form a compact Hausdorff topological semigroup with respect to matrix multiplication and M. Rosenblatt in [14] studied this problem for stochastic matrices by utilizing a previous theorem of his where he extended Kawada-Itó's result and gave a necessary and sufficient condition for the weak convergence of the convolution iterates μ^n of a probability measure μ on a compact semigroup. Although his necessary and sufficient condition is nice, there is still the difficulty of determining the smallest two-sided ideal of the semigroup generated by the support of the distribution of A_i, and Rosenblatt's theorem required first the determination of this ideal. The purpose of this paper is to extend Rosenblatt's work further in the special case of $d \times d$ bistochastic matrices as well as in the

0022-247X/91 $3.00

general noncompact case and the noncompact case of $d \times d$ nonnegative matrices.

We organize the paper as follows. In Section 2, we present a necessary and sufficient condition for the weak convergence of (μ^n) in a general locally compact semigroup. The condition here is in a simpler form than the one originally given by Rosenblatt in the compact case. In this section, we also present other results that are needed later and some results on nonnegative matrices.

In Section 3, we present some results on the structure of $d \times d$ bistochastic matrices. In Section 4, we utilize the results in Sections 2 and 3 to obtain our results on the weak convergence of μ^n in bistochastic matrices. In Section 5, we present *complete* solutions of the problem for the cases $d=2$, $d=3$, and $d=4$.

It is relevant to point out that definitions of terms such as convolution in a semigroup, a completely simple semigroup, etc., are all given in detail in [12], and therefore, are not repeated here.

2. Weak Convergence of μ^n in a Semigroup

First, we present a general theorem on the weak convergence of (μ^n).

THEOREM 2.1. *Let S be a locally compact second countable Hausdorff semigroup and μ be a probability measure on the Borel subsets of S. Suppose that*

$$S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n},$$

where S_μ is the support of μ.

Suppose that the sequence $\{\mu^n : n \geqslant 1\}$ is a tight sequence; that is, given $\varepsilon > 0$, there is a compact set K_ε such that for all $n \geqslant 1$, $\mu^n(K_\varepsilon) > 1 - \varepsilon$. Then the sequence $(1/n) \sum_{k=1}^n \mu^k$ converges weakly to a probability measure ν, where S_ν is the kernel K of S. The group factor G of K (which is completely simple) is compact. The sequence μ^n converges weakly to ν iff there does not exist a subgroup H of K such that the following conditions hold:

1. *H is a normal subgroup of the group $eKe \equiv G$, where e is the identity of H,*
2. *$YX \subset H$, where Y is the set of all idempotents in Ke and X is the set of all idempotent elements in eK,*
3. *$eS_\mu e \subset gH$ for some $g \in G \backslash H$.*

Furthermore, μ^n converges weakly iff $\liminf_n S_\mu^n$ *is nonempty, where we define* $\liminf_n S_\mu^n$ *as $\{x \in S :$ for every open set $V(x)$ containing x, there* *exists a positive integer N such that $n \geq N$ implies that $V(x) \cap S_\mu^n$ is nonempty$\}$.*

Proof. The first part and most of this theorem follows from Theorem 4.1 in [10]. It follows from there that $S_\nu = K$ and that the mapping Θ

$$\Theta: (x, g, y) \to x \cdot g \cdot y$$

from $X \times G \times Y$ to K, where $X \equiv$ the set of all idempotents in Ke, $G \equiv eKe$, and $Y \equiv$ the set of all idempotents in eK (e being any fixed idempotent in K and the set of all idempotents in K being nonempty), is an isomorphism if we define the multiplication in $X \times G \times Y$ as

$$(x, g, y)(x', g', y') = (x, g(yx')\, g', y').$$

Theorem 4.1 [10] also tells us that the sequence μ^n converges weakly to a probability measure ν iff there does not exist a normal subgroup H of G such that

$$S_\mu \cdot \Theta(X \times H \times Y) \subset \Theta(X \times gH \times Y) \tag{1}$$

with $YX \subset H$ for some $g \in G \backslash H$.

To finish the proof, we show that (1) is equivalent to the condition

$$eS_\mu e \subset gH. \tag{2}$$

To this end, suppose that (2) holds. Let $x \in S_\mu$ and $y = \Theta(e_1, h_1, f_1)$, where $h_1 \in H$, $e_1 \in X$, and $f_1 \in Y$. Note that $ex = e(ex) \in eK$ so that we can write

$$\Theta^{-1}(ex) = (e, h', f),$$

where $h' \in G$ and $f \in Y$.

We can also write

$$\begin{aligned}\Theta^{-1}(exe) &= \Theta^{-1}(ex)\, \Theta^{-1}(e)\\ &= (e, h', f)(e, e, e)\\ &= (e, h', e), \qquad \text{since } fe = e\end{aligned}$$

(Y is known to be a right-zero semigroup).

Thus, $exe = \Theta(e, h', e) = eh'e = h' \in gH$, by (2). Also,

$$\begin{aligned}\Theta^{-1}(exy) &= \Theta^{-1}(ex)\, \Theta^{-1}(y)\\ &= (e, h', f)(e_1, h_1, f_1)\\ &= (e, h'fe_1h_1, f_1)\\ &\in X \times gH \times Y,\end{aligned} \tag{3}$$

since $fe_1 \in H$ (recall : $YX \subset H$), $h_1 \in H$ and $h' \in gH$.

Now $xy \in K$ and we can write

$$\Theta^{-1}(xy) = (e_2, h_2, f_2), \qquad h_2 \in G. \tag{4}$$

Then we have

$$\begin{aligned}\Theta^{-1}(exy) &= \Theta^{-1}(e)\,\Theta^{-1}(xy)\\ &= (e, e, e)(e_2, h_2, f_2)\\ &= (e, h_2, f_2), \qquad \text{since } ee_2 = e\end{aligned}$$

(X being a left-zero semigroup), so that by (3),

$$h_2 = h'fe_1h_1 \in gH.$$

Thus, it follows by (4) that

$$xy \in X \times gH \times Y.$$

This means that (1) holds.

Now if we assume (1), then we have immediately

$$S_\mu \cdot (XHY) \subset X \cdot gH \cdot Y$$

so that

$$S_\mu \cdot e \subset X \cdot gH \cdot Y$$

or

$$\begin{aligned}eS_\mu e &= e(S_\mu e)\,e\\ &\subset (eX)\,gH(Ye)\\ &= egHe = gH\end{aligned}$$

implying (2).

Next we prove the last part of the theorem. Note that weak convergence of the sequence μ^n to the probability measure ν easily implies that $S_\nu \subset \liminf_n S_\mu^n$. To prove the converse, let us assume that $\liminf_n S_\mu^n$ is nonempty. Then it is easy to verify that $\liminf_n S_\mu^n$ is an ideal of S (because of the very definition of S). We establish the converse by contradiction.

Let us assume that (μ^n) does not converge weakly. Then by the first part of the theorem, it follows that there exists a normal subgroup H of the group G ($\equiv eKe$) such that for each positive integer N and some $g \in G \backslash H$,

$$eS_\mu^n e \subset g^n H, \tag{5}$$

where e and K are as before.

[Note that the inclusion above follows for each n since (1) implies also

$$\begin{aligned} &S_\mu^{n+1}\cdot\theta(X\times H\times Y)\\ &\quad\subset S_\mu\cdot\theta(X\times g^nH\times Y)\\ &\quad= S_\mu\cdot\theta(X\times H\times Y)\cdot\theta(X\times g^nH\times Y)\\ &\quad\subset\theta(X\times gH\times Y)\cdot\theta(X\times g^nH\times Y)\\ &\quad=\theta(X\times g^{n+1}H\times Y).] \end{aligned}$$

Let $x\in G$. Since G is a topological group, given any open set $U(x)$ containing x, there exist open sets $U(y)$ containing $y\in G$ and $U(z)$ containing $z\in G$ such that

$$U(x)\cap G\supset[eU(y)e]\cdot[eU(z)e]^{-1}, \tag{6}$$

where $[eU(z)e]^{-1}$ denotes the inversion in G.

Now K, being the kernel of S, is a subset of $\liminf_n S_\mu^n$ so that y and z are both elements of $\liminf_n S_\mu^n$. Hence, there exists a positive integer N such that

$$U(y)\cap S_\mu^N\neq\phi,\qquad U(z)\cap S_\mu^N\neq\phi. \tag{7}$$

It follows from (5) and (7) that

$$eU(y)e\cap g^NH\neq\phi,\qquad eU(z)e\cap g^NH\neq\phi.$$

Then this and (6) imply that

$$[U(x)\cap G]\cap H\neq\phi.$$

This means that $x\in H$. Thus, H cannot be a proper subgroup of G and this is a contradiction. Q.E.D.

We remark that Theorem 2.1 makes a very important point, namely that there will be convergence in distribution of products of i.i.d. $d\times d$ *real* or *complex* matrices if we know that the support of μ, the distribution of one of these, contains an idempotent element and the sequence (μ^n) is tight (since in this case, $\liminf_n S_\mu^n$ is nonempty).

The next three theorems are important and needed in the context of the present paper. We simply state them.

THEOREM 2.2 [9]. *Let S be a (multiplicative) semigroup of $d\times d$ bistochastic matrices closed in the usual topology. Then, it has a kernel K which is a finite group.*

THEOREM 2.3 [2]. *Let S be a semigroup of $d \times d$ real matrices such that S has a completely simple kernel K. Then, $K = m(S)$, where $m(S)$ is the set of all matrices in S with minimal rank.*

THEOREM 2.4 [8]. *Let μ and S be as in Theorem 2.1. Then the sequence μ^n is tight iff S has a completely simple kernel K with a compact group factor in its product representation such that for any open set $G \supset K$,*

$$\lim_{n \to \infty} \mu^n(G) = 1.$$

Below we show that we can be more specific (than in the general Theorem 2.1) for products of 3×3 and 2×2 random nonnegative matrices.

THEOREM 2.5. *Let μ be a probability measure on 3×3 nonnegative matrices (with usual topology and matrix multiplication). Let $S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n}$. Suppose that (μ^n) is tight and that S_μ is* not *contained in any finite group of invertible (rank 3) 3×3 nonnegative matrices. Then (μ^n) converges weakly iff there exists an idempotent e in S such that $e \in eS_\mu e$. (Note that when S_μ is contained in a finite group, easy necessary and sufficient conditions for the weak convergence of μ^n are well-known.)*

Proof. By the tightness of (μ^n), it follows from Theorem 2.1 that there is a completely simple kernel K of S and this kernel, by Theorem 2.3, consists of all matrices in S with minimal rank. If $K = \{0\}$, then, of course, $0 \in 0 \cdot S_\mu \cdot 0$ and by Theorem 2.1, μ^n converges weakly to the unit mass at the zero matrix. Suppose that the rank of the matrices in K is 1. In this case, for any idempotent e in K, eKe is a compact group of rank 1 nonnegative matrices, and, therefore, eKe must be a singleton, in particular, $eKe = \{e\}$ so that since $eS_\mu e \subset eKe$, $\{e\} = eS_\mu e$ and in this case, μ^n converges weakly by Theorem 2.1. Suppose that the rank of the matrices in K is 2 and that there exists an idempotent e_0 in S such that $e_0 \in e_0 S_\mu e_0$. (Note that we have only assumed e_0 in S. We do not know if this e_0 belongs to K.) Then $\operatorname{rank}(e_0) \geqslant 2$. If $\operatorname{rank}(e_0) = 3$, then e_0 is the identity matrix and then $S_\mu = e_0 S_\mu e_0$ contains the identity matrix. By the second part of Theorem 2.1, it follows that μ^n converges weakly. If $\operatorname{rank}(e_0) = 2$, then $e_0 \in K$ (by Theorem 2.3). Again it follows by Theorem 2.1 that μ^n converges weakly since the condition (3) in Theorem 2.1 cannot hold as $e_0 \in e_0 S_\mu e_0$. Finally, if the rank of the matrices in K is 2 and μ^n converges weakly, then for any idempotent e in K, the compact group eKe (being a compact group of nonnegative matrices) must be finite, in fact, eKe can have at most two elements (see [9]) so that $eS_\mu e$ is either $\{e\}$ or all of eKe. In either case, $e \in eS_\mu e$. Q.E.D.

Let us remark that a similar theorem holds for 2×2 nonnegative matrices.

Finally in this section we state without proof a general theorem on weak convergence and then give a necessary and sufficient condition for tightness in nonnegative matrices. (As far as we know, these theorems, as all the results in this section except for Theorem 2.2 through 2.4, are new.)

THEOREM 2.6. *Let S be a locally compact second countable semigroup, μ a Borel probability measure on S, and $S=\overline{\bigcup_{n=1}^{\infty} S_\mu^n}$. Suppose that (μ^n) is tight. Then S has a completely simple kernel K. Suppose that either S is abelian or K is compact. Then there exist elements $a_n \in K$ such that the sequence $\mu^n * \delta_{a_n}$ converges weakly to a probability measure ν such that $S_\nu \subset K$.*

(The proof of this theorem is not simple. It is omitted here since the theorem, although relevant in our context, is of a different character than the other results in this paper.)

Our last theorem in this section gives a necessary and sufficient condition for tightness of the sequence (μ^n) in $d\times d$ nonnegative matrices.

THEOREM 2.7. *Let μ be a probability measure on $d\times d$ nonnegative matrices (with usual topology and matrix multiplication). Let $S=\overline{\bigcup_{n=1}^{\infty} S_\mu^n}$. Let J be the set of all matrices in S, which have at least one zero row or one zero column or both, and $m(S)$ be the set of matrices in S with the minimal rank.*

Suppose that $m(S)\cap J^c \neq \phi$. Then the sequence (μ^n) is tight iff S is compact. (Note that $m(S)\cap J^c\neq\phi$ is a verifiable condition.)

Furthermore, if instead of assuming $m(S)\cap J^c\neq\phi$ we assume only that $O\notin S$, then the tightness of the sequence (μ^n) implies that S is nearly compact; that is, there is $M>0$ such that for each x in S, $\min_{i,j} x_{i,j}\leqslant M$. Conversely, if $\mu(J)=0$ and for some positive integer N,

$$\mu^N\{\textit{the strictly positive matrices in } S\}>0,$$

then near compactness of S implies the tightness of (μ^n).

Proof. We need to prove only the "only if" part. Suppose that (μ^n) is tight. By Theorem 2.1, $(1/n)\sum_{k=1}^{n}\mu^k$ converges weakly to a probability measure ν and S_ν is the completely simple kernel of S with a compact group factor. By Theorem 2.3, $S_\nu=m(S)$. Let $x\in m(S)\cap J^c$. Then $x\cdot m(S)\cdot x$ is a compact group of nonnegative matrices. Let e be the identity of this group. Then, $e\in J^c$ and we have $em(S)\,e=x\cdot m(S)\cdot x$. Since $m(S)$ is an ideal of S, $em(S)\,e=eSe$. The proof will be complete when we establish the following property:

(*) For any compact subset A of S, the set $Ax^{-1} = \{y \in S \mid yx \in A\}$ and $x^{-1}A = \{y \in S \mid xy \in A\}$ are compact when $x \in J^c \cap S$. (Note that $S = (e^{-1}[eSe])\, e^{-1}$.)

To prove (*), let $y \in Ax^{-1}$. Then, $yx \in A$. There is a positive δ such that $\sum_{j=1}^{d} x^{ij} \geqslant \delta > 0$, $1 \leqslant i \leqslant d$. Also by the compactness of A, there is a positive δ' (holds for all $y \in A$) such that

$$\delta y_{ij} \leqslant y_{ij} \cdot \sum_{k=1}^{d} x_{ik} \leqslant \sum_{k=1}^{d} (yx)_{ik} < \delta'$$

(for $1 \leqslant i, j \leqslant d$). This means that Ax^{-1} is bounded. Since Ax^{-1} is a closed subset of S, it is compact. Similarly, $x^{-1}A$ is compact.

For the second part of the theorem, suppose that $O \notin S$. Let A be a compact subset of S and $x \in S$. Then by the same method as above, there exist $M_1 > 0$, $M_2 > 0$ such that

$$Kx^{-1} \subset \{y \in S \mid \max_i \min_j y_{ij} \leqslant M_1\}$$

and

$$x^{-1}[Kx^{-1}] \subset \{y \in S \mid \min_{i,j} y_{ij} \leqslant M_2\}$$

The rest is simple. The converse part follows from [8] or [6]. Q.E.D.

We remark that when $m(S) \subset J$, S need not be compact in order that (μ^n) is tight. Examples demonstrating this are not difficult to find.

3. The Semigroup of $d \times d$ Bistochastic Matrices

Let B_d denote the set of all $d \times d$ bistochastic matrices. From now on, we denote the set $\{1, 2, ..., d\}$ by D.

Definition 3.1. By the base of an element x in B_d, we mean the finest partition $\{C_1, C_2, ..., C_r\}$ of D such that

(a) $x_{ij} = 0$ whenever $i \in C_\alpha$, $j \in C_\beta$, and $|C_\alpha| \neq |C_\beta|$;

(b) $\forall \alpha$, there is one and only one β such that $x|_{C_\alpha \times C_\beta}$ is a nonzero block;

(c) $\forall \beta$, there is one and only α such that $x|_{C_\alpha \times C_\beta}$ is a nonzero block.

Lemma 3.1. *Every element x in B_d has a unique base.*

Proof. Consider the set of all partition of D satisfying (a), (b), and (c) in Definition 3.1. This set is nonempty since the trivial partition $\{D\}$

belongs to it. Since this set is finite, it is clear that there exists a base. To prove its uniqueness, let us consider two partitions

$$\{X_1, X_2, ..., X_u\} \tag{8}$$

and

$$\{Y_1, Y_2, ..., Y_v\} \tag{9}$$

each satisfying (a), (b), and (c) above.

Consider the partition of all possible sets of the form $X_i \cap Y_j$ (called the refinement of (8) and (9)). Clearly (b) and (c) hold for this refinement. To establish (a), let

$$C_\alpha = X_a \cap Y_b, \qquad C_\beta = X_c \cap Y_d.$$

Suppose that $C_\alpha \times C_\beta$ is a nonzero block of x. Then

$$\sum_{i \in X_a} x_{ij} = 1 \qquad \forall j \in X_c, \tag{10}$$

$$\sum_{j \in X_c} x_{ij} = 1 \qquad \forall i \in X_a, \tag{11}$$

$$|X_a| = |X_c|, \qquad |Y_b| = |Y_d|. \tag{12}$$

Also, it is clear that

$$\begin{aligned} x_{ij} &= 0 \qquad \forall i \in C_\alpha, \qquad j \notin C_\beta; \\ &= 0 \qquad \forall i \notin C_\alpha, \qquad j \in C_\beta. \end{aligned} \tag{13}$$

It follows from (10), (11), and (13) that

$$\sum_{i \in X_a \setminus C_\alpha} x_{ij} = 1 \qquad \forall j \in X_c \setminus C_\beta,$$

$$\sum_{j \in X_c \setminus C_\beta} x_{ij} = 1 \qquad \forall i \in X_a \setminus C_\alpha.$$

Thus, $|X_c \setminus C_\beta| = |X_a \setminus C_\alpha|$. It follows from (12) that $|C_\alpha| = |C_\beta|$. Q.E.D.

LEMMA 3.2. *Let S be a compact semigroup in B_d. Let K be the kernel of S. Then K is a group and the base of its identity e is coarser than that of every element of S.*

Proof. By [9], K is a finite group and every element in K has the same base as that of e. In fact, if

$$\{C_1, C_2, ..., C_r\}$$

is the base of e, then given any y in K, there is a permutation $\pi(y)$ on the set $\{1, 2, ..., r\}$ such that $y|_{C_i \times C_j}$ is a strictly positive block or an all zero block according as $j = \pi(y)[i]$ or not.

Now let $y \in S$. Then $ye \in K$. Since K is a group, there is an element $z \in K$ such that $z \cdot ye = e$. But $z \cdot y \in K$ and $(zy)e = zy$. Thus, $zy = e$.

Let $u \in C_i$, $v \in C_j$. Then if $j \neq i$, $e_{uv} = 0$, and

$$0 = e_{uv} = \sum_{k \in C_{\pi(z)[i]}} z_{uk}\, y_{kv}.$$

Since for each $k \in C_{\pi(z)[i]}$, $z_{uk} > 0$, it is clear that the block

$$y|_{C_{\pi(z)[i]} \times C_j}$$

is a nonzero block iff $j = i$. This means that for the element y, the partition

$$\{C_1, C_2, ..., C_r\}$$

satisfies conditions (a), (b), and (c) in the definition of the base, so that the base of y can only be finer than that of e. Q.E.D.

In what follows, we write $x \approx y$ if x has the same nonzero block as y with respect to the base of y.

LEMMA 3.3. *Let e be an idempotent in B_d and let $x \in B_d$ have the same base as e. Let $L(x)$ be the set of all limit points of the sequence $\{x^n : n \geqslant 1\}$. Then $L(x) = \{e, ex, ..., ex^{k-1}\}$, where k is the smallest positive integer such that $x^k \approx e$, and $L(x)$ is a subgroup.*

Proof. It is clear that there is a smallest positive integer k such that $x^k \approx e$. Let us first show that $L(x)$ is a group, and as such finite. Clearly, $L(x)$ is an abelian semigroup. Also, for any $a \in L(x)$ and $b \in L(x)$, if

$$x^{n_i} \to a, \qquad x^{m_i} \to b,$$

then by compactness of B_d, there exist suitable subsequences $(p_i) \subset (n_i)$ and $(q_i) \subset (m_i)$ such that

$$x^{p_i - q_i} \to c$$

for some $c \in L(x)$; this means that

$$x^{p_i} = x^{p_i - q_i} \cdot x^{q_i} \to cb$$

or

$$a = cb.$$

Therefore, $L(x) \cdot b = L(x)$ for every $b \in L(x)$. Thus, $L(x)$ is an abelian (finite) group.

Next, we show that e is the identity of $L(x)$. To this end, let e' be the identity of $L(x)$. Then there is a subsequence p_i such that

$$x^{p_i k} \to e'.$$

Note that $L(x)$ is the kernel of the semigroup closure$(\{x^n : n \geqslant 1\})$ so that by Lemma 3.2, the base of e' is coarser than that of x. Thus, the base of e' is coarser than that of e. Since $x^{p_i k} \to e'$ and $x^k \approx c$, it is clear that e' and e have the same base. This implies, of course, that $e' = e$.

Now we claim that

$$x^{nk} \to e \qquad \text{as} \quad n \to \infty.$$

Suppose that $f \in L(x^k)$. Then, $fe = f$ (since e is the identity of $L(x^k) \subset L(x)$). But it is clear that $f \approx e$, since $x^k \approx e$. Thus, $fe = e$. Thus, $L(x^k) = \{e\}$. Since for $1 \leqslant s \leqslant k$,

$$x^{nk+s} \to e \cdot x^s,$$

$$L(x) = \{e, ex, ..., ex^{k-1}\}. \qquad \text{Q.E.D.}$$

In what follows, we say: $x \in P(e)$, where e is an idempotent in B_d, if $x \neq e$ and x belongs to a group with e as identity. Note that if

$$\{C_1, C_2, ..., C_r\}$$

is the base of e, then every permutation π on $\{1, 2, ..., r\}$ produces an element $x \in P(e)$, where for $1 \leqslant i \leqslant r$.

$$x|_{C_i \times C_j}$$

is a strictly positive block with each entry equals $|C_i|^{-1} = |C_j|^{-1}$ or an all zero block according to $j = \pi(i)$ or not. Every element in $P(e)$ occurs in this way.

Let I denote the set of all idempotents in B_d. Denote by P, the set $\bigcup \{P(e) : e \in I\}$.

Let us use the following notations; which will be used below and in later sections.

$$R = \{x \in B_d | \lim_{n \to \infty} x^n = e_0\},$$

where e_0 is the element in B_d with rank one.

$$S^* = B_d \backslash (R \cup P \cup I),$$

$$T(a) = \{x \in S^* | x \approx a\}$$

and

$$T(a)' = T(a) \cup \{a\}.$$

LEMMA 3.4. *Let* $C \subset B_d$ *such that* $C \subset \bigcup_{a \in P} T(a)'$ *and* $A = \{a \in P | C \cap T(a)' \neq \phi\}$. *Write*: $S = \overline{\bigcup_{n=1}^{\infty} C^n}$ *and* $S' = \overline{\bigcup_{n=1}^{\infty} A^n}$. *Then if* K *and* K' *are the kernels of* S *and* S' *respectively, then* $K = K'$.

Proof. Let us first show that $S' \subset S$. Let $a \in A$. Then there exists $x \in C$ such that either $x = a$ or else $x \in S^*$ and $x \approx a$. Let $e \sim a$ (that is, a and e have the same base), where e is an idempotent. Let $\{C_1, C_2, ..., C_r\}$ be the base of a as well as e. Then a corresponds to a permutation π on $\{1, 2, ..., r\}$ such that for $1 \leqslant i \leqslant r$, $a|_{C_i \times C_{\pi(i)}}$ is the nonzero block of a. It is easily verified that

$$xe = a = ex.$$

By Lemma 3.3, $e \in L(x)$ so that $a = ex$ also belongs to $L(x)$. Thus, $A \subset S$ or $S' \subset S$.

Now we claim that K', the kernel of S', is an ideal of S. To this end, we note that for any $y \in C$, there exists an $e \in I$ such that $y \sim e$ and $y \approx a$ for some $a \in P(e)$. By earlier arguments, $ye = a$. Let e' be the identity of K'. Then $yee' = ae'$. By Lemma 3.2, the base of e' is coarser than the base of any element in S' and therefore, it is coarser than the base of e since a above belongs to S' and $a \sim e$. This means that $ee' = e'$ so that $ye' = ae' \in K'$. Thus, $yK' = ye'K' \subset K'$ or $CK' \subset K'$ or $SK' \subset K'$.

Similarly, $K'S \subset K'$. This means that K' is an ideal of S and therefore, $K \subset K'$. Since K' is simple, $K = K'$. Q.E.D.

4. Weak Convergence of μ^n in B_d

Let μ be a probability measure in $d \times d$ bistochastic matrices with support S_μ. Let $S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n}$. Then S is the closed (in fact, compact) semigroup generated by S_μ where the topology in B_d is the usual topology. We define the nth convolution iterate as usual as

$$\mu^{n+1}(A) = \int \mu^n(Ax^{-1})\, \mu(dx),$$

where $Ax^{-1} = \{y \in S \mid yx \in A\}$. Weak convergence is defined as usual. See [12] for details.

Theorem 4.1. *Let K be the kernel of S. Then μ^n converges weakly iff there does not exist a proper normal subgroup H of the group K such that $S_\mu \cdot g \subset H$ for some $g \in K \backslash H$.*

Proof. Immediate from Theorem 2.1. Q.E.D.

Theorem 4.2. *Let μ and S be as before in B_d. Suppose that $S_\mu \cap (I \cup I') \neq \phi$, where I is the set of all idempotent elements in S and $I' = \{s \in S \mid s \approx e$ for some idempotent e in $S\}$. Then μ^n converges weakly.*

Proof. Immediate from Theorem 2.1. Q.E.D.

THEOREM 4.3. *Let μ and S be as in Theorem 4.2. If $S \cap R \neq \phi$, then μ^n converges weakly to the unit mass at e_0 (the rank 1 element in B_d). (Since R is dense in B_d, the same result holds when S has nonempty interior in B_d.)*

Proof. Note that if $x \in S \cap R$, then $e_0 \in S$ so that the kernel K of S is the singleton $\{e_0\}$. Since by Theorem 2.2, $\mu^n(G) \to 1$ as $n \to \infty$ for any open set $G \supset K$, the theorem follows. Q.E.D.

THEOREM 4.4. *Suppose that $S_\mu \subset \bigcup_{a \in P} \overline{T(a)}$. Let $A = \{a \in P;\ S_\mu \cap T(a)' = \phi\}$. Suppose that ν is a probability measure with support A. Then μ^n converges weakly iff ν^n converges weakly.*

Proof. Let $S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n}$, $S' = \overline{\bigcup_{n=1}^{\infty} A^n}$. By Lemma 3.4, both S and S' have the same kernel. Let e be the identity of this kernel. Then for any g in the kernel,

$$S_\mu \cdot g = S_\mu \cdot e \cdot g = S_\nu \cdot e \cdot g = S_\nu \cdot g$$

since for any $x \in S_\mu$, there exists $y \in S_\nu$ such that $x = ye$ and conversely, for any $y \in S_\nu$, there exists $x \in S_\mu$ such that $x = ye$ (look at the proof of Lemma 3.4). The theorem now follows by Theorem 4.1. Q.E.D.

We will apply these results in section 5 to solve the problem completely for $d=2$, $d=3$, and $d=4$.

5. WEAK CONVERGENCE OF μ^n IN 2×2, 3×3, AND 4×4 BISTOCHASTIC MATRICES

5.1. 2×2 *Bistochastic Matrices*

For a probability measure μ on 2×2 bistochastic matrices, if S_μ contains an element

$$\begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}, \qquad 0 < a < 1,$$

then μ^n converges weakly to the unit mass at $e_0 = \left(\begin{smallmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{smallmatrix}\right)$, since the kernel is $\{e_0\}$; otherwise, $S_\mu \subset \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \cup \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. In this case, μ^n converges weakly to the uniform distribution on this two-point group or the unit mass at the identity matrix, unless S_μ is a singleton other than the identity matrix.

5.2. 3×3 *Bistochastic Matrices*

Let μ be a probability measure on 3×3 bistochastic matrices with usual topology, with support S_μ.

Consider the following subsets:

$$A_1=\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 1-a \\ 0 & 1-a & a \end{pmatrix}; 0<a<1, a\neq 1/2\right\},$$

$$A_2=\left\{\begin{pmatrix} a & 0 & 1-a \\ 0 & 1 & 0 \\ 1-a & 0 & a \end{pmatrix}; 0<a<1, a\neq 1/2\right\},$$

and,

$$A_3=\left\{\begin{pmatrix} a & 1-a & 0 \\ 1-a & a & 0 \\ 0 & 0 & 1 \end{pmatrix}; 0<a<1, a\neq 1/2\right\}.$$

Let $\{e_1(\text{the identity matrix}), e_2, e_3, e_4, e_5, e_6\}$ be the six permutation matrices where each of $\{e_2, e_3, e_4\}$ has order 2 and each of $\{e_5, e_6\}$ has order 3.

Consider the three idempotent matrices (with rank 2) given by

$$f_1=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad f_2=\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \quad f_3=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $I=\{e_1, f_1, f_2, f_3\}$, $P=\{e_2, e_3, e_4, e_5, e_6\}$, $A=A_1\cup A_2\cup A_3$, and $R=(I\cup P\cup A)^c$. Then the following results hold:

(i) If $S_\mu\cap R\neq\phi$, then μ^n converges to the unit mass at e_0 (the bistochastic matrix with all the entries the same). In case $S_\mu\cap I\neq\phi$, then μ^n converges to the uniform distribution on a finite subgroup. If $S_\mu\cap[\{f_1, f_2, f_3\}\cup A]\neq\phi$, then the limit is the unit mass at one of the elements in $\{e_0, f_1, f_2, f_3\}$.

When the limit exists and is different from the unit mass at one of the elements in $\{e_0, f_1, f_2, f_3\}$, then S_μ must be contained in a finite subgroup and the limit is the uniform distribution on the smallest such subgroup.

(ii) Suppose that $S_\mu\subset P$. When S_μ is either a singleton or contained in $\{e_2, e_3, e_4\}$, then μ^n does not converge; otherwise, μ^n converges and the limit is the uniform distribution on the subgroup $\{e_1, e_2, e_3, e_4, e_5, e_6\}$, unless $S_\mu=\{e_5, e_6\}$ in which case the limit is the uniform distribution on the subgroup $\{e_1, e_2, e_3\}$.

Let us now prove these results. Let $S^* = A_1 \cup A_2 \cup A_3$, and $R = B_3 \setminus (I \cup P \cup S^*)$. Suppose that $S_\mu \cap S^* \neq \phi$. If there exist $x, y \in S_\mu$ such that $x \in A_i$, $y \in A_j$ $(i \neq j)$, then assuming $i = 1$ and $j = 2$, by Lemma 3.2, the base of every element in the kernel K of $S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n}$ is coarser than the partition $\{\{1\}, \{2, 3\}\}$ (the base of the elements in A_1) and $\{\{1, 3\}, \{2\}\}$ (the base of elements in A_2) so that the base of the elements in K is $\{\{1, 2, 3\}\}$. This means that $K = \{e_0\}$, the element in B_3 with rank one. It is then clear that μ^n converges weakly to the unit mass at e_0.

If $S_\mu \subset A_1$, then the base of the elements in K is $\{\{1\}, \{2, 3\}\}$. This means that $K = \{f_1\}$ so that μ^n converges weakly to the unit mass at f_1.

If $S_\mu \cap (R \cup I) \neq \phi$, then μ^n converges weakly, by Theorems 4.2 and 4.3.

Suppose now that $S_\mu \subset P$.

In case $S_\mu = \{e_i\}$, $i = 2, 3$ or 4, the kernel K of S is $\{e_1, e_i\}$ and Theorem 4.1 tells us immediately that μ^n does not converge. The same results hold when $S_\mu = \{e_i\}$, $i = 5$ or 6, in which case $K = \{e_1, e_5, e_6\}$ and the only proper normal subgroup is $\{e_1\}$; thus μ^n does not converge.

If $S_\mu \subset \{e_2, e_3, e_4\}$, but S_μ is not a singleton, then $K = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ and $S_\mu \cdot e_i \subset \{e_1, e_5, e_6\}$, a proper normal subgroup of K, for $i = 2$, 3 or 4; therefore, μ^n does not converge.

When $S_\mu = \{e_5, e_6\}$, $K = \{e_1, e_5, e_6\}$, the only (proper) normal subgroup is $\{e_1\}$ and $S_\mu \cdot e_5 = \{e_1, e_6\}$; therefore μ^n converges.

Finally, suppose that S_μ contains at least one of $\{e_5, e_6\}$ and at least one of $\{e_2, e_3, e_4\}$, then $K = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, with two proper normal subgroups $\{e_1, e_5, e_6\}$ and $\{e_1\}$. It follows easily that μ^n converges.

5.3. 4×4 *Bistochastic Matrices*

In this section, I' denotes the set $\{s \in S : s \approx e$ for some idempotent e in $S\}$, where $x \approx y$ means that x has the same nonzero block as y with respect to the base of y. (For definitions of R, I, P, S^* see the text preceding Lemma 3.4.)

Consider the following possibilities:

1. $S_\mu \cap (R \cup I \cup I') \neq \phi$.

By Theorems 4.2 and 4.3, μ^n converges in this case.

2. $S_\mu \subset (R \cup I \cup I')^c = \bigcup_{a \in P} T(a)'$

(a) $S_\mu \not\subset S_4$ and $S_\mu \cap T(a)' \neq \phi$, $S_\mu \cap T(b)' \neq \phi$

for some a, b in P $(a \neq b)$, where S_4 is the group of 4×4 permutation matrices.

In this case, μ^n converges.

(b) $S_\mu \subset T(a)'$ for some $a \in P$.

In this case, μ^n does *not* converge.

3. $S_\mu \subset S_4$.

In this case, one can use the classical Kawada-Itó theorem.

Proof of Case 2(a). Note that in what follows, the set P contains the elements f_i', g_j', x_k, y_t, z_m, where $1 \leqslant i, m \leqslant 6$, $1 \leqslant k \leqslant 9$, $1 \leqslant j \leqslant 3$, $1 \leqslant t \leqslant 8$, and the set I contains the elements e_i, f_j, g_k, e, where $0 \leqslant i \leqslant 4$, $1 \leqslant j \leqslant 6$, $1 \leqslant k \leqslant 3$. All these elements have been described in the Appendix.

Let $A = \{a \in P \mid S_\mu \cap T(a)' \neq \phi\}$, Then $A_0 = \{a, b\} \subset A$. Let ν be a probability measure with support A. We prove that ν^n converges weakly and then use Theorem 4.4.

First, we determine K, the kernel of $S' = \overline{\bigcup_{n=1}^{\infty} A^n}$.

(i) $K = \{e_0\}$.

This happens when A_0 is none of the following:

$$\{f_1', g_1'\}, \{f_6', g_1'\}, \{f_3', g_2'\}, \{f_5', g_2'\}, \{f_2', g_3'\}, \text{ or } \{f_4', g_3'\}; \tag{14}$$

$$\{\{f_i', f_j'\} : i \neq j\} \backslash \{\{f_1', f_6'\}, \{f_2', f_5'\}, \{f_3', f_4'\}\}; \tag{15}$$

$$\{f_1', f_6'\}, \qquad \{f_2', f_5'\}, \qquad \text{or} \quad \{f_3', f_4'\}; \tag{16}$$

$$\{f_j', x_i\}, \qquad \text{where } j = 1 \text{ or } 6 \text{ and } i = 1, 4 \text{ or } 5; \tag{17}$$

$$\text{or} \quad j = 2 \text{ or } 4 \text{ and } i = 2, 8 \text{ or } 9;$$

$$\text{or} \quad j = 3 \text{ or } 5 \text{ and } i = 3, 6 \text{ or } 7.$$

$$\{g_j', x_i\}, \qquad \text{where } j = 1 \text{ and } i = 1, 4 \text{ or } 5; \tag{18}$$

$$\text{or} \quad j = 2 \text{ and } i = 2, 8 \text{ or } 9;$$

$$\text{or} \quad j = 3 \text{ and } i = 3, 6 \text{ or } 7.$$

Clearly in this case, μ^n converges to the mass at e_0.

(ii) $K = \{g_i, g_i'\}$, for $i = 1, 2$ or 3.

This happens if A_0 is one of the subsets in (14). In this case, the only proper normal subgroup is $H = \{g_i\}$ and $f_i' g_j' = g_j' \notin H$, $1 \leqslant i \leqslant 6$. Hence, ν^n (and therefore, μ^n) converges.

(iii) $K = \{e_i\}$, for $i = 1, 2, 3$ or 4.

This happens when A_0 is one of the elements in (15). In this case, μ^n converges to the unit mass at e_i.

(iv) $K = \{g_i\}$, for $i = 1, 2$ or 3.

This happens when A_0 is one of the sets in (16). In this case μ^n also converges.

(v) When A_0 is one of the sets in (17), then $S' = \{f_j', f_j, x_i, e\}$ and $K = \{f_j, f_j'\}$. The only proper normal subgroup of K is then $\{f_j\}$ and $x_i f_j' = f_j' \notin H$.

Thus, ν^n converges weakly and therefore, so does μ^n.

(vi) When A_0 is one of the sets in (18), then $S' = \{g_j', g_j, x_i, e\}$ and $K = \{g_j, g_j'\}$. In this case, the only proper normal subgroup of K is $H = \{g_j\}$ and $x_i g_j' = g_j' \notin H$.

Thus, ν^n converges; consequently, μ^n converges.

Case 2(b). $S_\mu \subset T(a)'$ and $a \notin S_4$.
Let $A = \{a\}$ and $\nu(A) = 1$. Then

$$\nu^{2n}(A) \to 0 \qquad \text{and} \qquad \nu^{2n+1}(A) \to 1 \qquad \text{as } n \to \infty.$$

Hence, ν^n does not converge and therefore, μ^n also does not. Q.E.D.

Appendix: Classification of 4 × 4 Bistochastic Matrices

Notations

(a) $B_d = \{x; x \text{ is a } 4 \times 4 \text{ bistochastic matrix.}\}$

(b) $e_0 = \begin{pmatrix} \frac14 & \frac14 & \frac14 & \frac14 \\ \frac14 & \frac14 & \frac14 & \frac14 \\ \frac14 & \frac14 & \frac14 & \frac14 \\ \frac14 & \frac14 & \frac14 & \frac14 \end{pmatrix}$

(c) $e_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac13 & \frac13 & \frac13 \\ 0 & \frac13 & \frac13 & \frac13 \\ 0 & \frac13 & \frac13 & \frac13 \end{pmatrix}, \quad e_2 = \begin{pmatrix} \frac13 & 0 & \frac13 & \frac13 \\ 0 & 1 & 0 & 0 \\ \frac13 & 0 & \frac13 & \frac13 \\ \frac13 & 0 & \frac13 & \frac13 \end{pmatrix}$

$e_3 = \begin{pmatrix} \frac13 & \frac13 & 0 & \frac13 \\ \frac13 & \frac13 & 0 & \frac13 \\ 0 & 0 & 1 & 0 \\ \frac13 & \frac13 & 0 & \frac13 \end{pmatrix}, \quad e_4 = \begin{pmatrix} \frac13 & \frac13 & \frac13 & 0 \\ \frac13 & \frac13 & \frac13 & 0 \\ \frac13 & \frac13 & \frac13 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$f_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac12 & \frac12 \\ 0 & 0 & \frac12 & \frac12 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac12 & \frac12 & 0 \\ 0 & \frac12 & \frac12 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

$f_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac12 & 0 & \frac12 \\ 0 & 0 & 1 & 0 \\ 0 & \frac12 & 0 & \frac12 \end{pmatrix}, \quad f_4 = \begin{pmatrix} \frac12 & 0 & 0 & \frac12 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac12 & 0 & 0 & \frac12 \end{pmatrix}$

$$f_5=\begin{pmatrix}\frac12&0&\frac12&0\\0&1&0&0\\\frac12&0&\frac12&0\\0&0&0&1\end{pmatrix},\qquad f_6=\begin{pmatrix}\frac12&\frac12&0&0\\\frac12&\frac12&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

$$g_1=\begin{pmatrix}\frac12&\frac12&0&0\\\frac12&\frac12&0&0\\0&0&\frac12&\frac12\\0&0&\frac12&\frac12\end{pmatrix},\qquad g_2=\begin{pmatrix}\frac12&0&\frac12&0\\0&\frac12&0&\frac12\\\frac12&0&\frac12&0\\0&\frac12&0&\frac12\end{pmatrix}$$

$$g_3=\begin{pmatrix}\frac12&0&0&\frac12\\0&\frac12&\frac12&0\\0&\frac12&\frac12&0\\\frac12&0&0&\frac12\end{pmatrix},\qquad e=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

(d) $$f_1'=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&\frac12&\frac12\\0&0&\frac12&\frac12\end{pmatrix},\qquad f_2'=\begin{pmatrix}0&0&0&1\\0&\frac12&\frac12&0\\0&\frac12&\frac12&0\\1&0&0&0\end{pmatrix}$$

$$f_3'=\begin{pmatrix}0&0&1&0\\0&\frac12&0&\frac12\\1&0&0&0\\0&\frac12&0&\frac12\end{pmatrix},\qquad f_4'=\begin{pmatrix}\frac12&0&0&\frac12\\0&0&1&0\\0&1&0&0\\\frac12&0&0&\frac12\end{pmatrix}$$

$$f_5'=\begin{pmatrix}\frac12&0&\frac12&0\\0&0&0&1\\\frac12&0&\frac12&0\\0&1&0&0\end{pmatrix},\qquad f_6'=\begin{pmatrix}\frac12&\frac12&0&0\\\frac12&\frac12&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}$$

$$g_1'=\begin{pmatrix}0&0&\frac12&\frac12\\0&0&\frac12&\frac12\\\frac12&\frac12&0&0\\\frac12&\frac12&0&0\end{pmatrix},\qquad g_2'=\begin{pmatrix}0&\frac12&0&\frac12\\\frac12&0&\frac12&0\\0&\frac12&0&\frac12\\\frac12&0&\frac12&0\end{pmatrix}$$

$$g_3'=\begin{pmatrix}0&\frac12&\frac12&0\\\frac12&0&0&\frac12\\\frac12&0&0&\frac12\\0&\frac12&\frac12&0\end{pmatrix}$$

$$x_1=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad x_2=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix},\quad x_3=\begin{pmatrix}1&0&0&1\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{pmatrix}$$

$$x_4=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\quad x_5=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix},\quad x_6=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$$

$$x_7=\begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{pmatrix},\quad x_8=\begin{pmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{pmatrix},\quad x_9=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}$$

$$y_1=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\\0&1&0&0\end{pmatrix},\quad y_2=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix},\quad y_3=\begin{pmatrix}0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$$

$$y_4=\begin{pmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1\end{pmatrix},\quad y_5=\begin{pmatrix}0&0&1&0\\0&1&0&0\\0&0&0&1\\1&0&0&0\end{pmatrix},\quad y_6=\begin{pmatrix}0&0&0&1\\0&1&0&0\\1&0&0&0\\0&0&1&0\end{pmatrix}$$

$$y_7=\begin{pmatrix}0&1&0&0\\0&0&0&1\\0&0&1&0\\1&0&0&0\end{pmatrix},\quad y_8=\begin{pmatrix}0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0\end{pmatrix},$$

$$z_1=\begin{pmatrix}0&0&1&0\\0&0&0&1\\0&1&0&0\\1&0&0&0\end{pmatrix},\quad z_2=\begin{pmatrix}0&0&0&1\\0&0&1&0\\1&0&0&0\\0&1&0&0\end{pmatrix},\quad z_3=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}$$

$$z_4=\begin{pmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\\0&0&1&0\end{pmatrix},\quad z_5=\begin{pmatrix}0&1&0&0\\0&0&0&1\\1&0&0&0\\0&0&1&0\end{pmatrix},\quad z_6=\begin{pmatrix}0&0&1&0\\1&0&0&0\\0&0&0&1\\0&1&0&0\end{pmatrix}$$

(e) $\mathscr{A}_1 = \{\{1, 2, 3, 4\}\}$

$\mathscr{B}_1 = \{\{1\}, \{2, 3, 4\}\}$

$\mathscr{B}_2 = \{\{2\}, \{1, 3, 4\}\}$

$\mathscr{B}_3 = \{\{3\}, \{1, 2, 4\}\}$

$\mathscr{B}_4 = \{\{4\}, \{1, 2, 3\}\}$

$\mathscr{C}_1 = \{\{1, 2\}, \{3, 4\}\}$

$\mathscr{C}_2 = \{\{1, 3\}, \{2, 4\}\}$

$\mathscr{C}_3 = \{\{1, 4\}, \{2, 3\}\}$

$\mathscr{D}_1 = \{\{1\}, \{2\}, \{3, 4\}\}$

$\mathscr{D}_2 = \{\{1\}, \{3\}, \{2, 4\}\}$

$\mathscr{D}_3 = \{\{1\}, \{4\}, \{2, 3\}\}$

$\mathscr{D}_4 = \{\{2\}, \{3\}, \{1, 4\}\}$

$\mathscr{D}_5 = \{\{2\}, \{4\}, \{1, 3\}\}$

$\mathscr{D}_6 = \{\{3\}, \{4\}, \{1, 2\}\}$

$\mathscr{E}_1 = \{\{1\}, \{2\}, \{3\}, \{4\}\}$

References

1. P. Bougerol, Tightness of products of random matrices and stability of linear stochastic systems, *Ann. Probab.* **15** (1987), 40–74.
2. W. E. Clark, Remarks on the kernel of a matrix semigroup, *Czechoslovak Math. J.* **15** No. 90 (1987), 305–309.
3. D. R. Cox and V. Isham, "Point Processes," Chapman & Hall, New York, 1980.
4. G. Högnäs, A note on products of random matrices, *Statist. Probab. Lett.* **5** 367–370.
5. G. Högnäs, Random semigroup acts on a finite set, *J. Austral. Math. Soc. Ser. A* **23** 481–498.
6. H. Kesten and F. Spitzer, Convergence in distribution of products of random matrices, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **67** (1984), 363–386.
7. V. M. Maximov, Convergence of non-homogeneous bistochastic Markov chains, *Theory Probab. Appl.* **15** (1970), 604–618.
8. A. Mukherjea, Convergence in distribution of products of random matrices : A semigroup approach, *Trans. Amer. Math. Soc.* **303** (1987), 395–411.
9. A. Mukherjea, Completely simple semigroup of matrices, *Semigroup Forum* 3 (1986), 405–429.
10. A. Mukherjea, Limit theorem : Stochastic matrices, ergodic Markov chains and measures on semigroups, *in* "Probability Analysis and Related Topics" (A. T. Bharucha-Reid, Ed.), Vol. 2, pp. 143–203, Academic Press, New York, 1979.

11. A. Mukherjea, Tightness of products of i.i.d. random matrices, *Probab. Th. Rel. Fields* **87** (1991), 389–401.
12. A. Mukherjea and N. A. Tserpes, "Measures on Topological Semigroups: Convolution Products and Randon Walks," Lecture Notes in Mathematics," Vol. 547, Springer-Verlag, New York, 1976.
13. M. Rosenblatt, "Markov Processes: Structure and Asymptotic Behavior," Springer-Verlag, New York, 1971.
14. M. Rosenblatt, Products of independent and identically distributed random stochastic matrices, *J. Math. Anal. Appl.* **11** (1965), 1–10.